Bidirectional Long Short-Term Memory Model for Metro Passenger flow Prediction

Gu Huang [1], Guitang Mai [2,*] and Tianxi Xiao [3]

[1] School of Traffic and Transportation Engineering, Central South University Changsha, Hunan, 410075, China

[2] School of international Education, Neusoft Institute Guangdong, Foshan, Guangdong, 528200, China

[3] International Education College, Don River College, Shandong Jiaotong University, Shandong, 250000, China

* Corresponding Author Email: Guitang2.Mai@live.uwe.ac.uk

Abstract. To maximize metro operations, accurate short-term passenger flow projections are essential. The paper utilizes Bidirectional Long Short-Term Memory (BiLSTM) to forecast passenger flow based on data from January 16 to January 25, 2019. By contrasting BiLSTM with conventional models of Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN), the research highlights BiLSTM's superior ability to capture temporal dependencies from both past and future data. The analysis reveals distinct patterns for weekdays and weekends, with double peaks during commute hours on weekdays and a continuous peak in the afternoon on weekends. The results indicate that BiLSTM outstandingly enhances the prediction's accuracy, with less Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) in contrast to RNN and LSTM. This enhanced predictive capability supports more effective scheduling and flow management in metro systems, ensuring better service and operational efficiency. The study underscores the practical benefits of BiLSTM in handling complex, dynamic passenger flow data.

Keywords: Metro; LSTM; BiLSTM; passenger flow prediction.

1. Introduction

This study's objective is to use the BiLSTM algorithm to accurately predict subway passenger flow. By analyzing historical data and external factors, a high-precision prediction model is built to optimize subway train scheduling and passenger flow control strategies so as to raise the subway system's quality of service and operational efficiency. This study will verify the applicability and reliability of this model for different time periods and passenger flow changes and prove its effectiveness in practical applications. Currently, methods for traffic flow prediction mainly include deep learning, machine learning, and statistical techniques [1]. In terms of statistical techniques, Liu et al. used the Autoregressive Integrated Moving Average Model (AIMAM) for predictive research on rail traffic flow [2]. Wu et al. employed ensemble approaches for reinforcement learning and time series decomposition to predict urban rail traffic flow [3]. Regarding machine learning methods, Li et al. combined decision tree methods with time series models, developing a new hybrid decision tree for transitory bus passenger flow prediction [4]. Che and Chen used vector machines to build an experimental platform for analyzing and predicting urban rail traffic passenger flow, achieving improvements over traditional prediction methods [5].

In terms of deep learning methods, Vinayakumar et al. used various RNN networks and found that the LSTM method was more accurate for traffic network flow prediction compared to other methods [6]. Yang et al. used Transformer and residual networks for short-term forecasting of passenger traffic across multiple transportation forms [7]. Liu employed an enhanced deep belief network for predicting temporary traffic flow [8]. The integration of multiple influence factors into the BiLSTM model significantly enhances the prediction performance for short-term passenger flow, making it a valuable tool for real-time operational management of rail transit systems [9]. The study highlighted the superiority of LSTM-based models over traditional statistical methods, demonstrating improved accuracy in daily passenger flow predictions, which is crucial for daily operational planning and management [10].

Although the above methods enhance the forecast accuracy to a certain extent, there are still some shortcomings in navigating the subway's intricate and diverse passenger flow. BiLSTM, as an improved LSTM model, can more successfully depict the data's front and back dependencies and provide more accurate prediction results. Therefore, this study selects BiLSTM to analyze historical passenger flow data and related external factors and build a high-precision passenger flow prediction model.

2. Methods

2.1. Data Source and Description

The sources of the data utilized in this study comes from the 24-hour card swiping information set of metro passengers in a certain city from January 16, 2019 to January 25, 2019, and the data are accurate and representative. The data field types in the dataset are shown in Table 1.

Table 1. Subway passengers in a certain city

Column name	Type	Explain	Example
time	String	Swiping time	2019/1/16 0:00
lineID	String	ID of metro line	C
stationID	Int	ID of station	64
deviceID	Int	ID of device	2980
status	Int	Entry and exit status, 0 for exit, 1 for entry	0
userID	String	ID of user	Bee069dae5399509d4427e1bda7a344ff
payType	Int	Swiping type	0

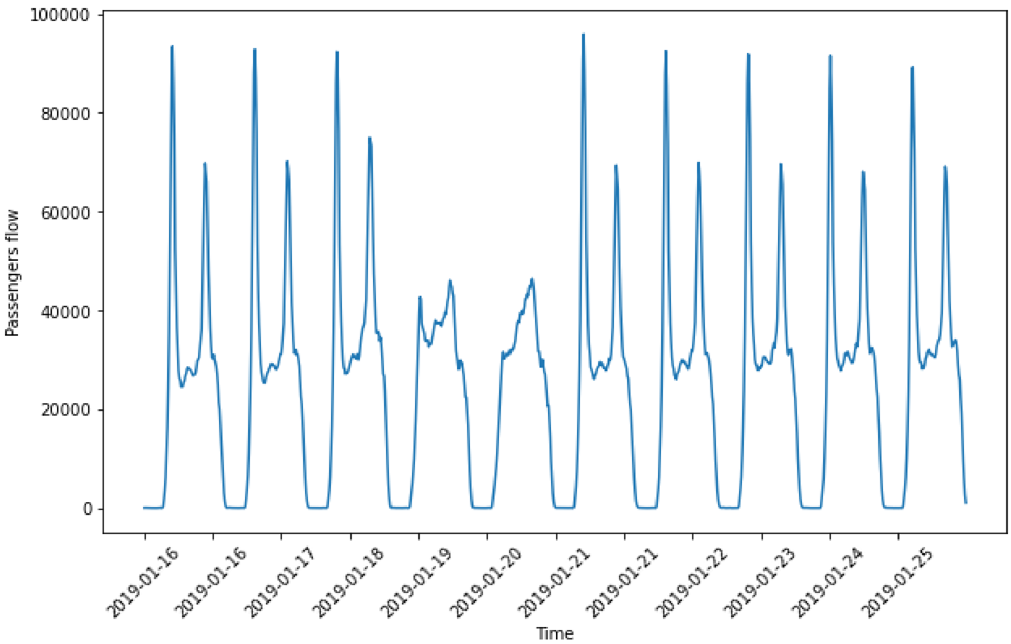

Fig. 1 Time sequence diagram of data on the flow of passengers arriving

This study mainly explores the rules and forecast of short-term flow of passengers within the metro. Besides, the study chooses the flow of passengers arriving as the standard of passenger flow. Therefore, this paper selects the data with a status of 1 in the original data set, removes indicators other than time, records the passenger flow within each time interval, and sets the time granularity to 15 minutes to obtain a new data set. All the data were plotted as a time series diagram, as demonstrated in Figure 1.

2.2. Method Introduction

2.2.1. LSTM model

LSTM is an improved cyclic network based on RNN. LSTM introduces forgotten door, output gate and input gate based on RNN so as to resolve the problem of prolonged reliance on RNN. The schematic illusion of LSTM structure is demonstrated in Figure 2.

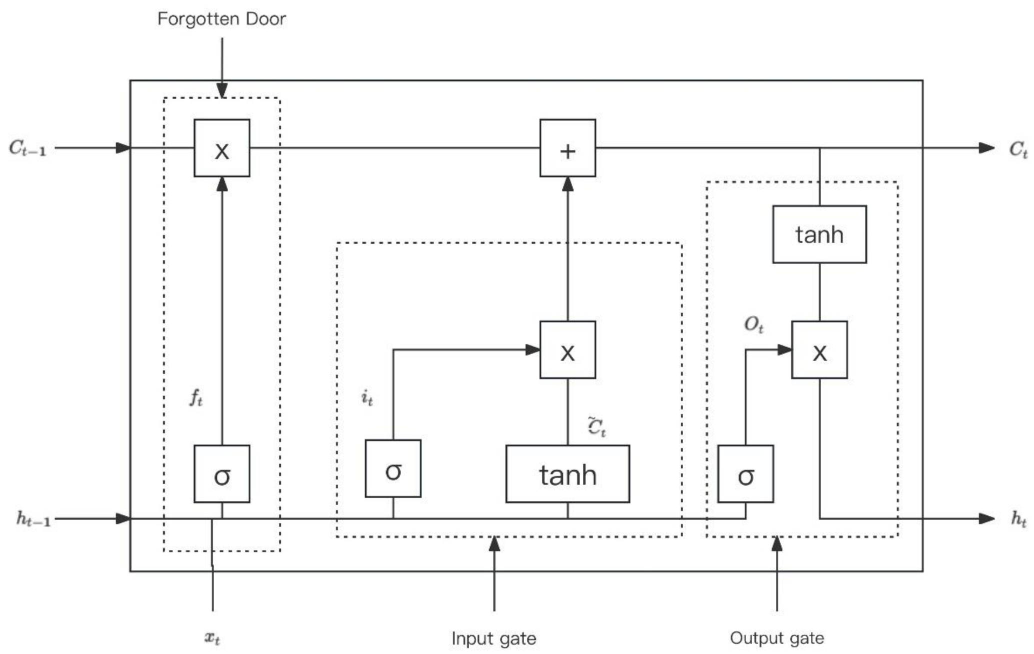

Fig. 2 The schematic diagram of LSTM structure

LSTM is an effective and commonly used forecasting method that can avoid the phenomenon of gradient disappearing and the problem of long-term dependence.

2.2.2. BiLSTM model

BiLSTM is an improved method on the LSTM network. Unlike the traditional LSTM network, in which information can only be transmitted from the past to the future, the BiLSTM network contains two LSTM networks. The two LSTM networks are independent of each other and process the input sequences in chronological and reverse chronological order, respectively. The schematic illustration of the structure of BiLSTM is demonstrated in Figure 3.

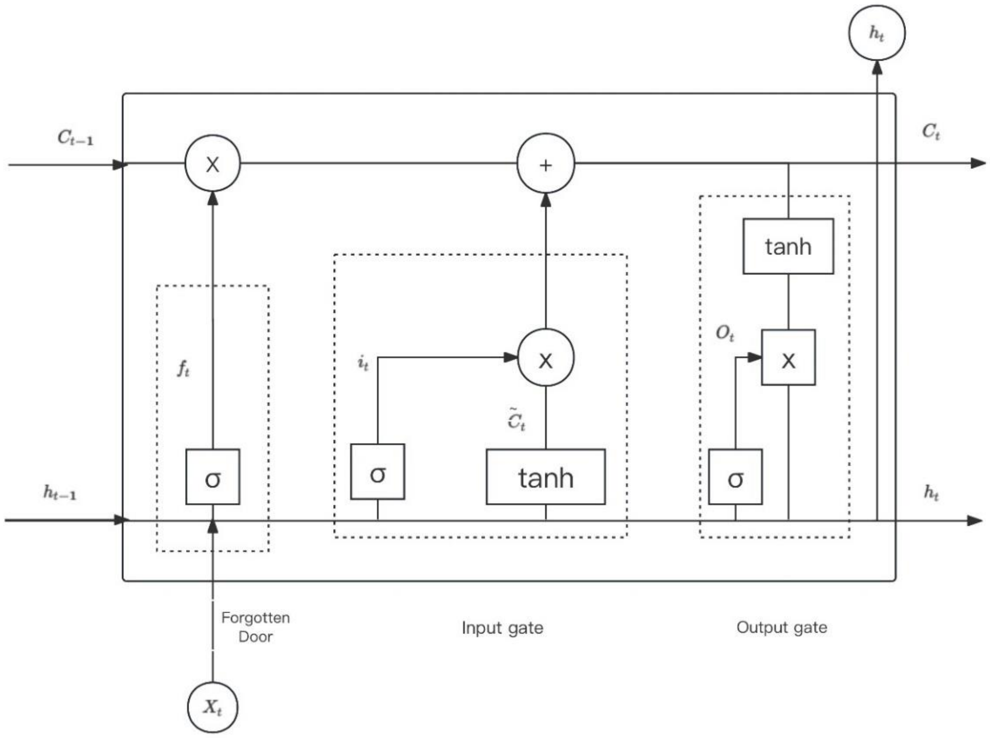

Fig. 3 The schematic illustration of the structure of BiLSTM

BiLSTM can make use of both past and future information to forecast, which makes it more accurate than the LSTM network to forecast, so it is ideal for short-term passenger flow projection in metros.

3. Results and Discussion

3.1. Descriptive Analysis

Through the preliminary analysis of Figure 1, this paper believes that the subway passenger flow presents different passenger flow trends both on the weekdays and on the weekends, respectively. To verify this conjecture, the Pearson correlation coefficient was introduced in this paper to examine the patterns of traveler flow on several days. The analysis results are shown in Figure 4.

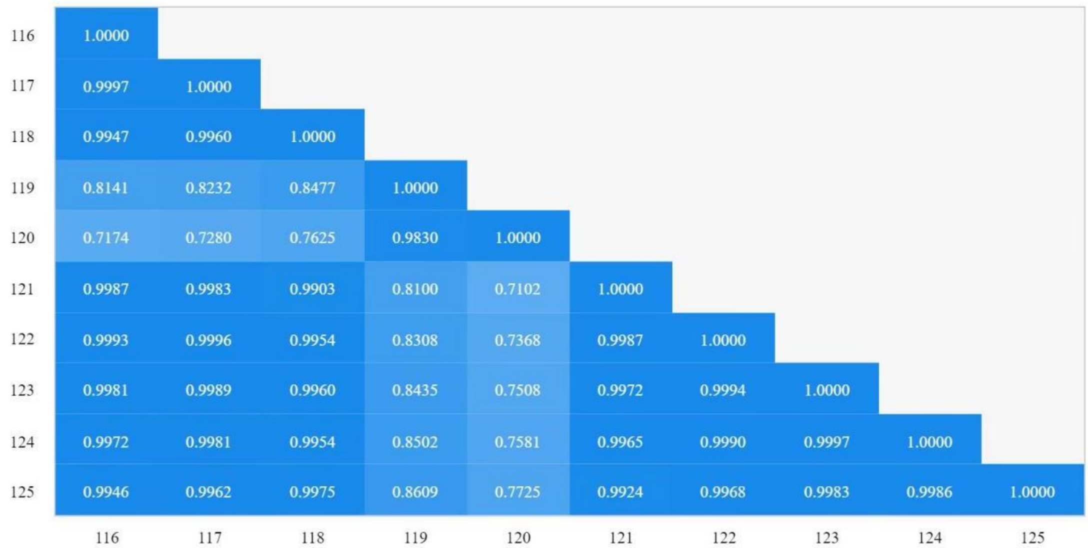

Fig. 4 Correlation of ridership trends from January 16 to January 25

In this paper, it is assumed that the data exhibit a robust association. when the Pearson coefficient is greater than 0.9. As can be seen from Figure 4, the person coefficient between January 19 and January 20 (i.e., 119 and 120 in the figure) at the weekend is greater than 0.9, showing a strong correlation; the person coefficient between the data at other times, i.e., working days, is greater than 0.9, showing a strong correlation. Therefore, this paper divides subway passenger flow development modes into two types: the first is weekday mode, as shown in Figure 5, and the second is weekend mode, as shown in Figure 6.

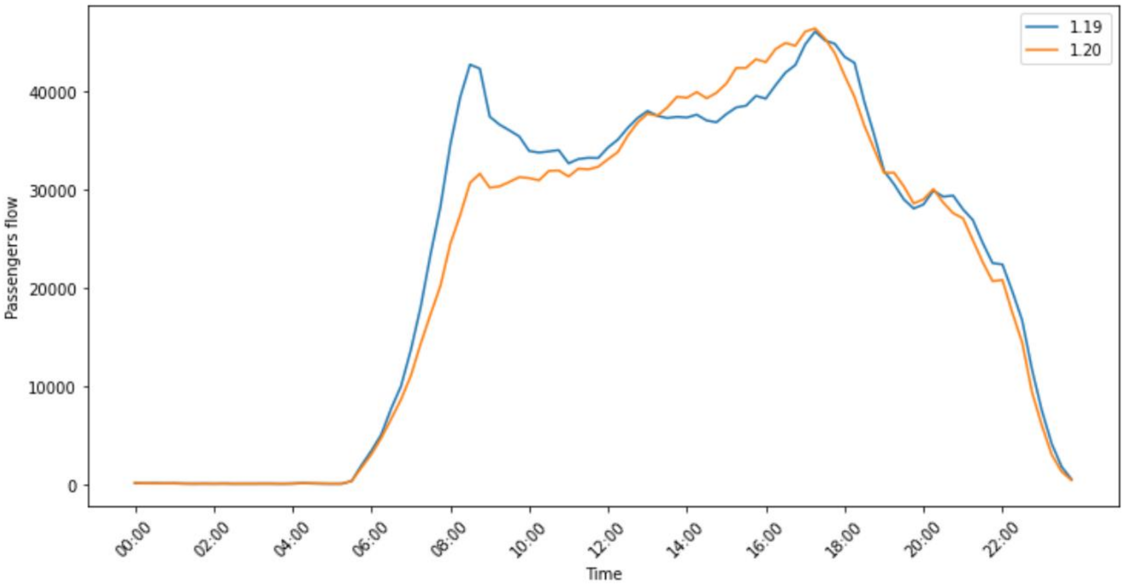

Fig. 5 Daily passenger flow inbound volume in working day mode

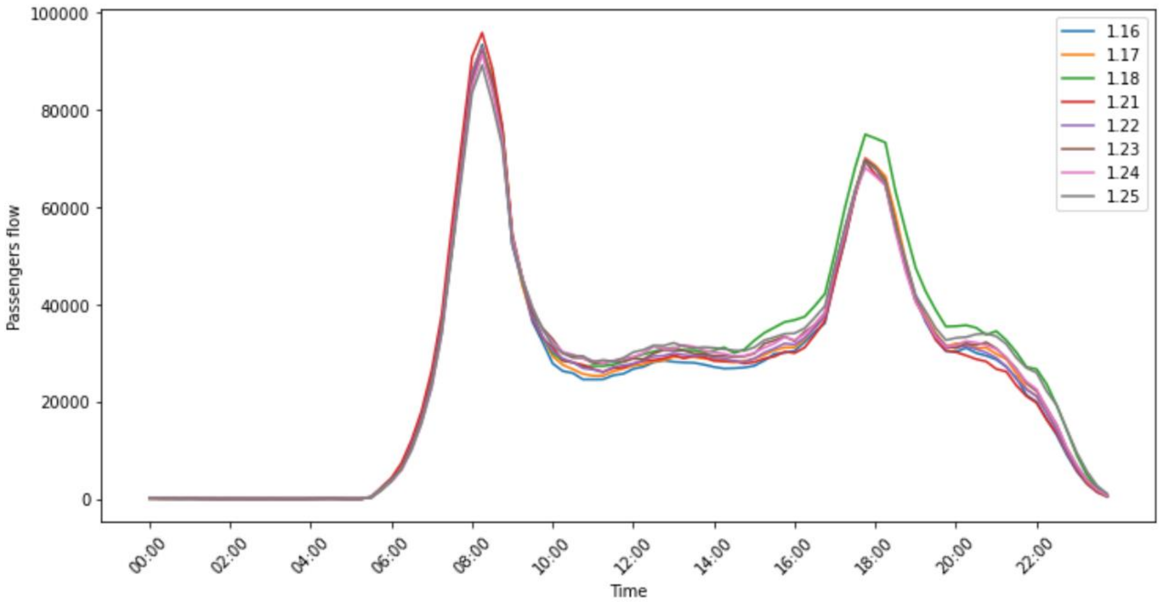

Fig. 6 Daily passenger flow inbound volume in weekend mode

In this paper, the two development models are further analyzed according to the two trends. As for the daily passenger flow inbound volume in weekday mode, it can be seen from the figure that there is a double peak in the morning and evening, which accords with the actual situation of increasing subway passenger flow due to morning and evening commutes in real life. As for the inbound daily passenger flow in weekend mode, it can be seen from the figure that there is a continuous peak phenomenon in the afternoon, which accords with the actual situation of the increase in afternoon subway passenger flow caused by weekend passengers going out to play in real life. To sum up, the division of passenger flow trends in this paper accords with the phenomenon of real life, which has strong practical significance.

3.2. BiLSTM Model Results

In this paper, linear normalization is used to put the data into the closed interval [0, 1], and then the ratio of the data sample division is 8:1:1. The BiLSTM parameters used in the study are demonstrated in Table 2.

Table 2. BiLSTM Parameter Settings

Time steps	neurons	Activation function	Optimizer	Loss function	Batch size	Epochs	Learning rate
10	50	Relu	Adam	Mean Squared Error (MSE)	32	100	0.001

The loss function diagram for 100 iterations is shown in Figure 7. It can be seen from the image that the iteration starts in a basic convergence state at about 20 times, and the validation loss curve is closer to 0 than the train loss curve after convergence, indicating that the fitting results have generalization.

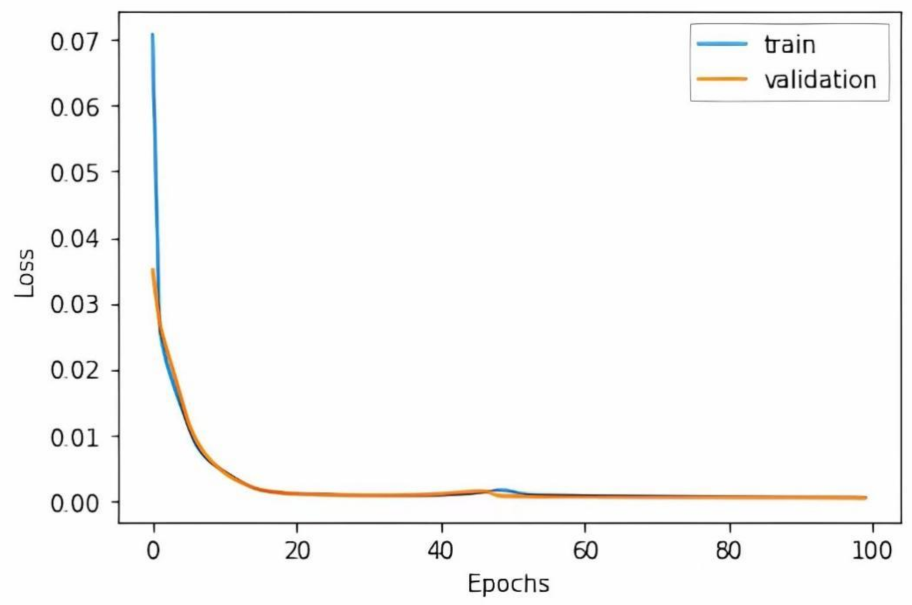

Fig. 7 BiLSTM function loss function diagram

The raw and predicted data are represented in the same time series diagram, as shown in Figure 8. The predicted curve and the actual curve basically coincide, indicating that the fitting effect is excellent.

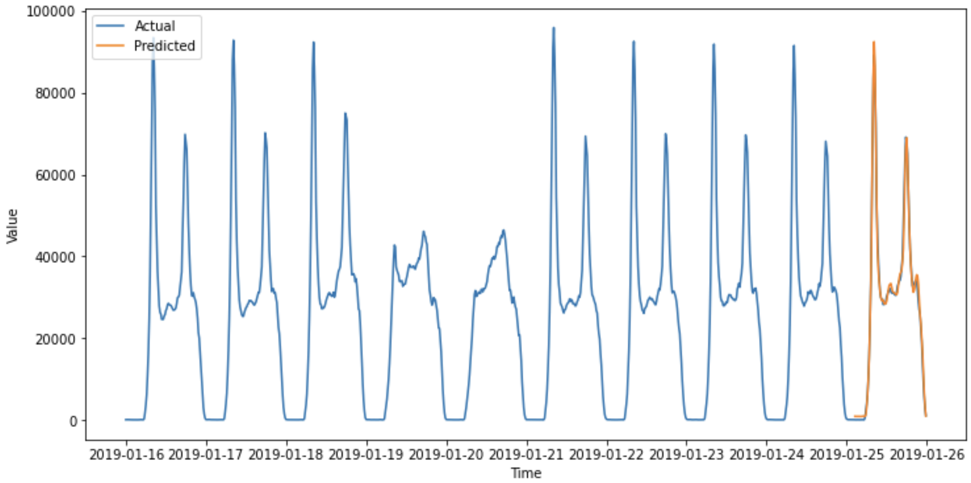

Fig. 8 Timing diagram of raw data and test data

3.3. Comparison Results

By using RMSE and MAE, the model's fitting effect was quantified. The above indexes of BiLSTM were contrasted with those of LSTM and RNN, and Table 3 was obtained.

Table 3. Comparison of RMSE and MAE indexes between BiLSTM model and other models

Prediction Model	RMSE	MAE
RNN	0.022	0.017
LSTM	0.020	0.014
BiLSTM	0.018	0.012

It can be seen from Table 3 that the BiLSTM has a better fitting effect than the RNN and the LSTM, indicating that the predicted data obtained by the BiLSTM model is closer to the true value.

4. Conclusion

To reliably forecast metro passenger flow, BiLSTM is employed in this study, aiming at optimizing subway train scheduling and passenger flow control strategies and improving the subway system's level of service and operational efficiency. The data were collected from subway card swipe records from January 16 to January 25, 2019, and inbound passenger flow was selected as the research object, and a high-precision BiLSTM prediction model was constructed. By analyzing the passenger flow trends in different time periods, the differences between weekday and weekend passenger flow patterns are verified. The findings indicate that BiLSTM is higher than LSTM and RNN in terms of RMSE and MAE, can better record the connection between before and after data of time series, and improves prediction accuracy. This study proves the advantage of the BiLSTM model in processing complex and changeable subway passenger flow data and provides important technical support for real-time operation management of the subway system.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Fang H, et al. Metro Station functional clustering and dual-view recurrent graph convolutional network for metro passenger flow prediction. Expert Systems with Applications, 2024, 247: 122550.

[2] Liu et al. Research on forecast of rail traffic flow based on ARIMA model. Journal of Physics: Conference Series, 2021, 1792(1): 12065.

[3] Wu P, et al. A Time Series Decomposition and Reinforcement Learning Ensemble Method for Short-Term Passenger Flow Prediction in Urban Rail Transit. Urban Rail Transit, 2023, 9(4): 323-351.

[4] Li P, Wu W, Pei X. A separate modelling approach for short-term bus passenger flow prediction based on behavioural patterns: A hybrid decision tree method. Physica A: Statistical Mechanics and its Applications, 2023, 616: 128567.

[5] Che S, Chen Y. Rail transit passenger volume forecasting method based on support vector machine. Third International Conference on Green Communication, Network, and Internet of Things (CNIoT 2023). SPIE, 2023, 12814: 500-505.

[6] Vinayakumar R, Soman K P, Poornachandran P. Applying deep learning approaches for network traffic prediction. 2017 International Conference on Advances in Computing, Communications, and Informatics (ICACCI). IEEE, 2017: 2353-2358.

[7] Yang Y, et al. Short-term passenger flow prediction for multi-traffic modes: A Transformer and residual network based multi-task learning method. Information Sciences, 2023, 642: 119144.

[8] Yong Liu. Short-Term Traffic Flow Prediction with Structure-Optimized Deep Belief Network. Sustainability in Environment, 2022.

[9] Qi Q, Cheng R, Ge H. Short-term inbound rail transit passenger flow prediction based on BILSTM model and influence factor analysis. Digital Transportation and Safety, 2023, 2(1): 12-22.

[10] Zhu H, Yang X, Wang Y. Prediction of daily entrance and exit passenger flow of rail transit stations by deep learning method. Journal of Advanced Transportation, 2018, 1: 6142724.